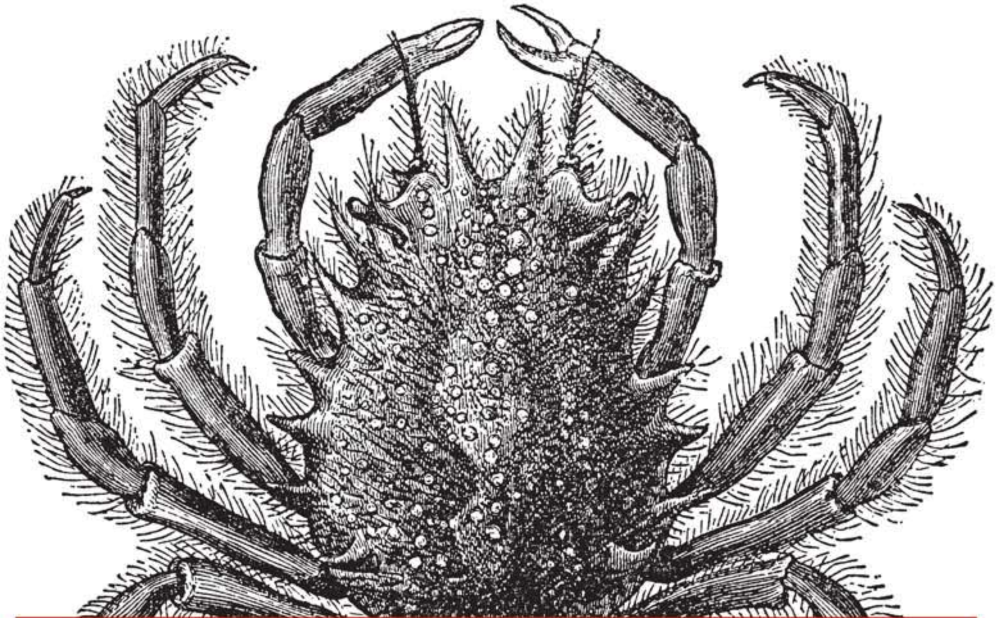
EXHIBIT J

OMNIBUS BROWN DECLARATION



STATISTICS IN A NUTSHELL

A Desktop Quick Reference

Sarah Boslaugh
& Paul Andrew Watters

O'REILLY®

Math/Statistics

STATISTICS IN A NUTSHELL

Need to learn statistics for your job? Want help passing that statistics course? *Statistics in a Nutshell* is a clear and concise introduction and reference for anyone who's new to the subject. This book gives you a solid understanding of statistics without the numbing complexity of most textbooks.

Each chapter presents easy-to-follow descriptions illustrated by graphics, formulas, and plenty of solved examples. Before you know it, you'll be applying statistical reasoning and techniques to everything from basic concepts of probability and hypothesis testing to multivariate analysis. *Statistics in a Nutshell* includes:

- Introductory material—Learn basic concepts of measurement and probability theory, data management for statistical analysis, research and experimental design, and how to critique statistics presented by others
- Basic inferential statistics—Discover the concepts of correlation, dichotomous and categorical data, and the distinction between parametric and nonparametric statistics
- Advanced inferential techniques—Learn the General Linear Model, ANOVA and MANOVA, and multiple linear regression
- Specialized techniques—Use and interpret business and quality improvement statistics, medical and public health statistics, and educational and psychological statistics

If you need to know how to perform most common statistical analyses—and how to use a wide range of statistical techniques without getting in over your head—this is the book for you.

Sarah Boslaugh, Ph.D., has been a statistical analyst for 15 years and currently teaches Intermediate Statistics at Washington University Medical School in St. Louis.

Paul Andrew Watters, Ph.D., CITP, is an associate professor in the School of Information and Mathematical Sciences and Centre for Informatics and Applied Optimization (CIAO) at the University of Ballarat in Victoria, Australia.

O'REILLY®
www.oreilly.com

US $34.99 CAN $34.99
ISBN: 978-0-596-51049-7

53499
9 780596 510497

Statistics in a Nutshell
by Sarah Boslaugh and Paul Andrew Watters

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*safari.oreilly.com*). For more information, contact our corporate/institutional sales department: (800) 998-9938 or *corporate@oreilly.com*.

Editor: Mary Treseler
Production Editor: Sumita Mukherji
Copyeditor: Colleen Gorman
Proofreader: Emily Quill
Indexer: John Bickelhaupt
Cover Designer: Karen Montgomery
Interior Designer: David Futato
Illustrator: Robert Romano

Printing History:

July 2008: First Edition.

RepKover. This book uses RepKover™, a durable and flexible lay-flat binding.

ISBN: 978-0-596-51049-7
[M]

H_0 is called the null hypothesis: in this example, the null hypothesis states that drug X is no improvement over standard treatment. H_A, sometimes written as H_1, is called the alternative hypothesis: in this case, the alternative hypothesis is that drug X is more effective than standard treatment. Note that the null and alternative hypotheses must be both mutually exclusive (no results could satisfy both conditions) and exhaustive (all possible results will satisfy one of the two conditions).

In this example, the alternative hypothesis is *single-tailed*: we state that the blood pressure of the group treated with drug X must be lower than that of the standard treatment group for the null hypothesis to be rejected. We could also state a *two-tailed* alternative hypothesis if that were more appropriate to our research question. If we were interested in whether the blood pressure of patients treated with drug A was different, either higher or lower, than that of patients receiving standard treatment, we would state this using a two-tailed alternative hypothesis:

H_0: $\mu_1 = \mu_2$
H_A: $\mu_1 \neq \mu_2$

Normally the first two steps would be performed before the experiment is designed or the data collected; the statistic to be used for hypothesis testing is also sometimes specified at this time, or is implicit in the hypothesis and type of data involved. We then collect the data and perform the statistical calculations, in this case probably a *t*-test or ANOVA, and based on our results make one of two decisions:

- Reject the null hypothesis and accept the alternative hypothesis, or
- Fail to reject the null hypothesis

The first case is sometimes called "finding significance" or "finding significant results." The process of statistical testing involves establishing a probability level or *p*-value (a topic treated in greater detail below) beyond which we will consider results from our sample strong enough to support rejection of the null hypothesis. In practice, the *p*-value is commonly set at 0.05. Why this particular value? It's an arbitrary cutoff point and dates back to the early twentieth century, when statistics were computed by hand and the results compared to published tables used to determine whether a result was significant or not. The use of $p < 0.05$ as the standard for significant results has been challenged (see the upcoming sidebar, "Controversies About Hypothesis Testing") but still remains common practice for published research. Alternative lower values are sometimes used, such as $p < 0.01$ or $p < 0.001$, but no one has been successful in legitimizing the use of a higher cutoff, such as $p < 0.10$.

Note that failure to reject the null hypothesis does not mean that we have proven it to be true, only that the experiment or study did not find sufficient evidence to reject it.

Inferential statistics allows us to make probabilistic statements about the data, but the possibility of error is inherent in the process. Statisticians have classified two types of errors when making decisions in inferential statistics, and set levels for error rates that are commonly considered acceptable. The two types of error are displayed in Table 7-1.

Controversies About Hypothesis Testing

Despite the ubiquitousness of hypothesis testing in modern statistical practice, and the canonical place that the $\alpha = 0.05$ significance level seems to have achieved, many researchers have criticized both practices. One of the chief critics is Jacob Cohen, whose arguments are presented in, among other places, his 1994 article "The World Is Round ($p < 0.05$)." (*American Psychologist*, 49:2, December 1994, 997–1003). Valid as many of these criticisms are, both hypothesis testing and the 0.05 significance level don't seem to be going away anytime soon. And some level must be set at which results are considered significant, to avoid attributing significance to differences due to chance or to sampling error. A sensible compromise is to realize that there's nothing magical about 0.05, even if it is sometimes treated as such, and that the significance level of a sample calculation is affected by many factors, including the size of the sample involved. It's a common saying among statisticians that if you have a large enough sample, any little difference will be statistically significant. The take-home message is that statistical methods don't relieve the researcher of the need to apply common sense.

Table 7-1. Type I and Type II errors

		True state	
		H_0 true	H_A true
Decision based on sample statistic	Accept H_0	Correct decision: H_0 true and H_0 not rejected	Type II error or β
	Reject H_0	Type I error or α	Correct decision: H_0 false and H_0 rejected

The diagonal boxes represent correct decisions: H_0 is true and was not rejected in the study, or H_0 is false and was rejected in the study. The other two boxes (often referred to as the off-diagonal boxes) represent Type I and Type II errors. A Type I error, also known as alpha or α, represents the error made when the null hypothesis is true but is rejected in a study. A Type II error, also called beta or β, represents the error made when H_0 is false but is not rejected in a study.

The level of acceptability for Type I error is conventionally set at 0.05, as noted above. Setting alpha at 0.05 means that we accept the fact of a 5% probability of Type I error. To put it another way, we understand when setting the alpha level at 0.05 that we accept the fact that in our study we have a 5% chance of rejecting the null hypothesis when we should have accepted it.

Type II error has received less attention in statistical theory because historically it has been considered a less serious error to fail to make an inference that is true (Type II error) than to make an inference that is false (Type I error). Conventional levels of acceptability for Type II error are $\beta = 0.1$ or $\beta = 0.2$. If $\beta = 0.1$, that means the study has a 10% probability of a Type II error, i.e., there is a 10% chance that the null hypothesis will be false but will be accepted in the study.